Computation for the Introduction to MCMC Chapter of *Handbook of Markov Chain Monte Carlo*

By
Charles J. Geyer
Technical Report No. 679
School of Statistics
University of Minnesota
July 29, 2010

The Monte Carlo standard error, expressed relative to the standard deviation of the invariant distribution is

```
> sqrt((1 + rho)/(1 - rho)/nsim)
```

```
[1] 0.1410674
```

In order to get a 10% relative error, we would need sample size

```
> ceiling((1 + rho)/(1 - rho)/0.1^2)
```

```
[1] 19900
```

In order to get a 10% relative error, we would need sample size

```
> ceiling((1 + rho)/(1 - rho)/0.01^2)
```

```
[1] 1990000
```

Variance inflation factor (VIF)

```
> (1 + rho)/(1 - rho)
```

```
[1] 199
```

Effective sample size

```
> nsim/((1 + rho)/(1 - rho))
```

```
[1] 50.25126
```

Figure 1 (page 3) is made by the following R statements

```
> par(mar = c(5, 4, 1, 1) + 0.1)
> plot(out, ylab = "x")
```

Figure 2 (page 4) is made by the following R statements

```
> foo <- cumsum(out)/seq(along = out)
> par(mar = c(5, 4, 1, 1) + 0.1)
> plot(as.ts(foo), ylab = expression(hat(mu)[n]), xlab = expression(n))
```

Figure 3 (page 5) is made by the following R statements

```
> lag.max <- 500
> par(mar = c(5, 4, 1, 1) + 0.1)
> plot(acf(out, lag.max = lag.max), ci.col = "black",
+     main = "")
> curve(rho^x, from = 0, to = lag.max, add = TRUE,
+     lty = 3, lwd = 2)
```

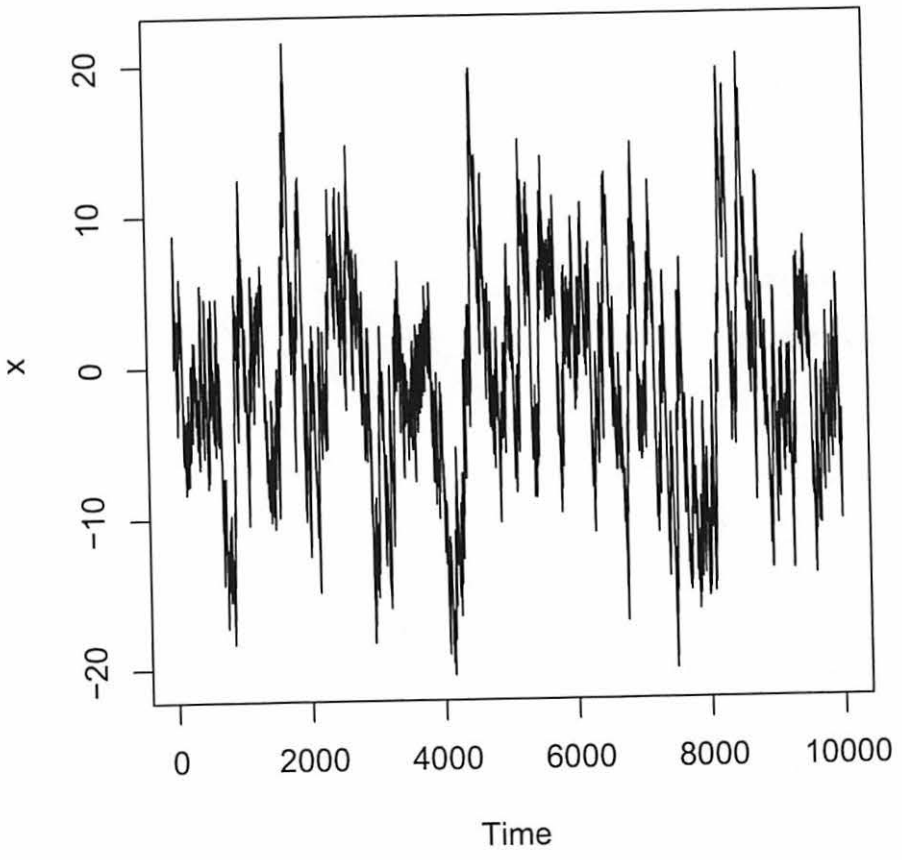

Figure 1: Time series plot of AR(1) time series.

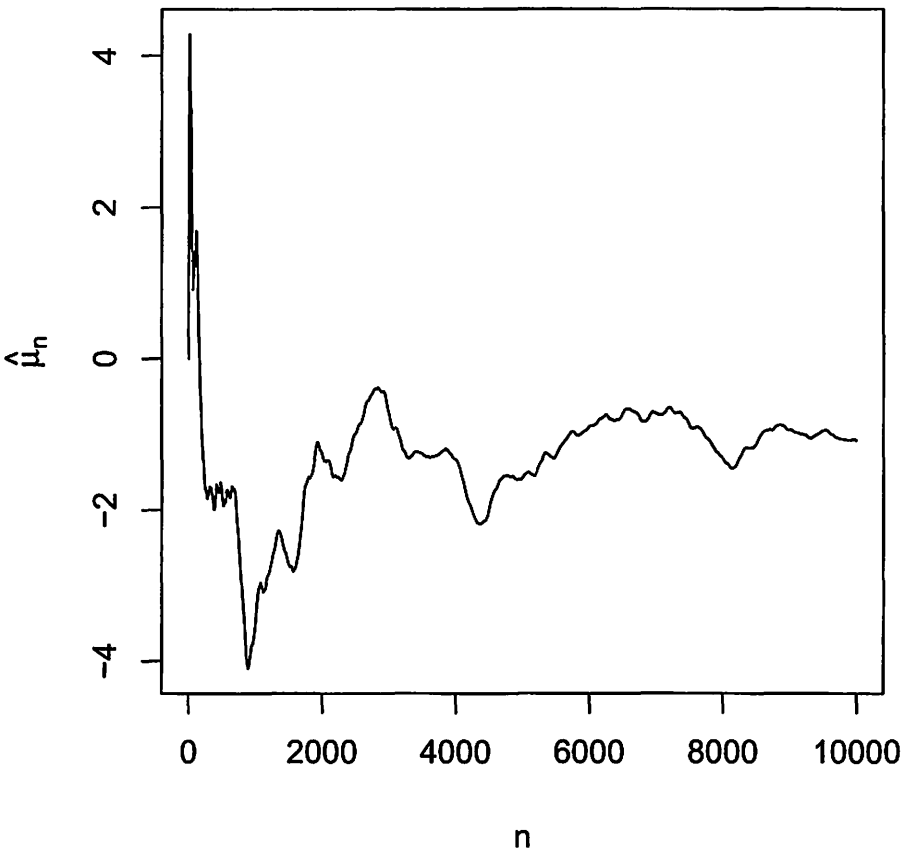

Figure 2: Sample means of AR(1) time series.

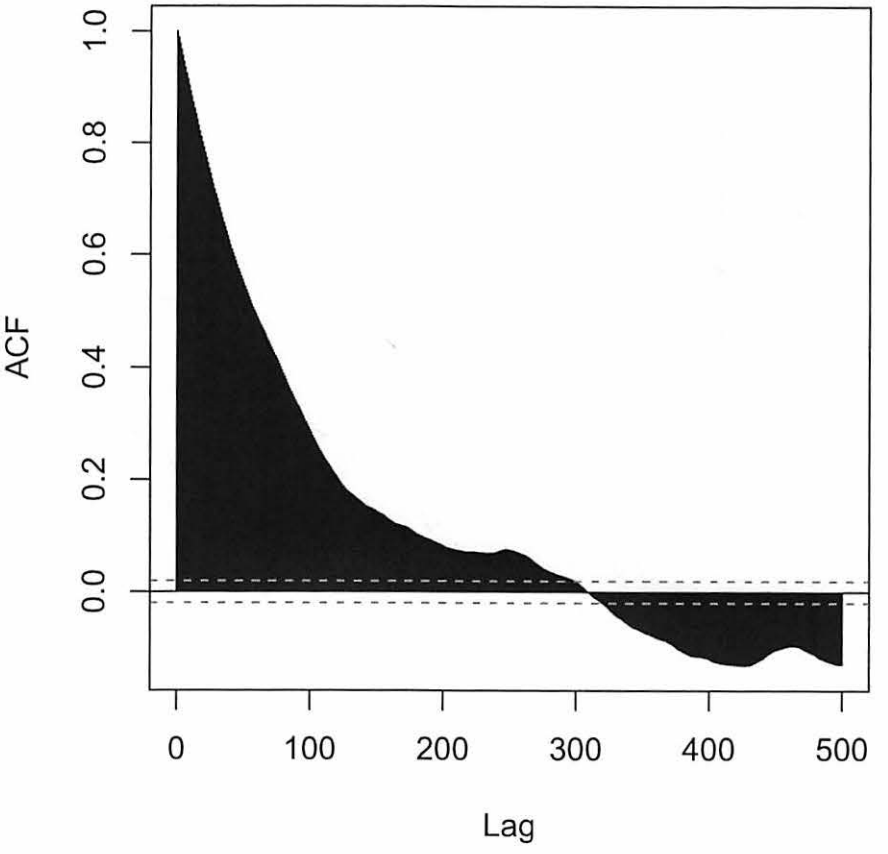

Figure 3: Autocorrelation plot of AR(1) time series. Dashed lines: 95% confidence intervals assuming white noise input. Dotted line: simulation truth autocorrelation function.

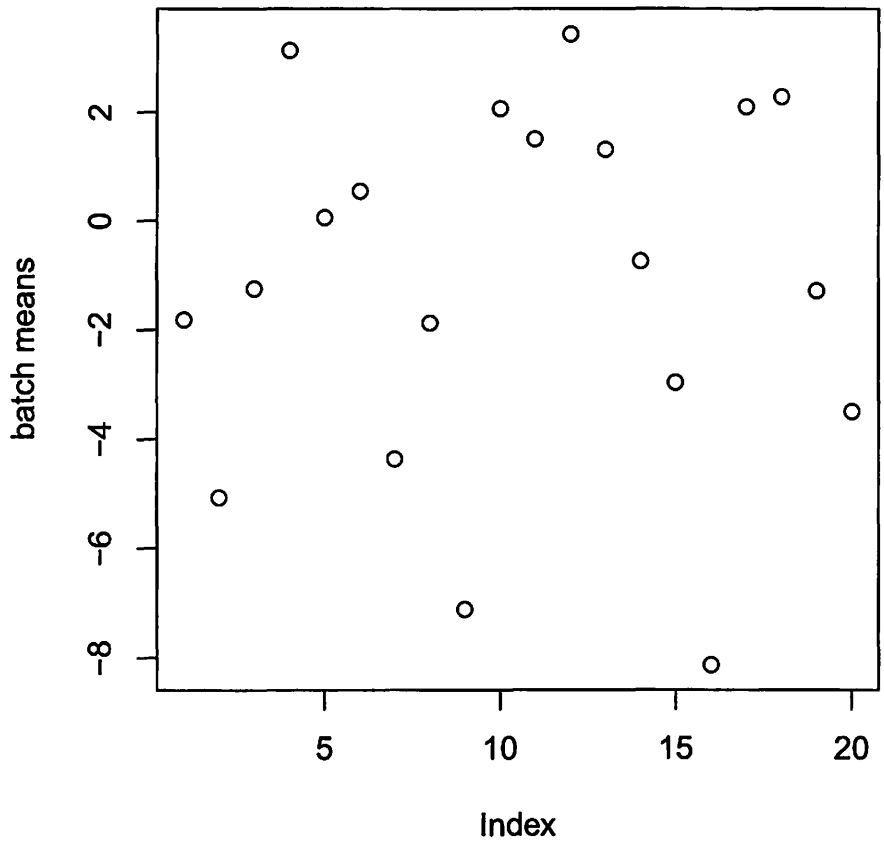

Figure 4: Batch mean plot of AR(1) time series.

2.1 Nonoverlapping Batch Means

Figure 4 (page 6) is made by the following R statements

```
> blen <- 500
> batch <- matrix(out, nrow = blen)
> batch <- apply(batch, 2, mean)
> par(mar = c(5, 4, 1, 1) + 0.1)
> plot(batch, ylab = "batch means")
```

Figure 5 (page 7) is made by the following R statements

```
> par(mar = c(5, 4, 1, 1) + 0.1)
> plot(acf(batch), ci.col = "black", main = "")
```

```
> t.test(batch)
```

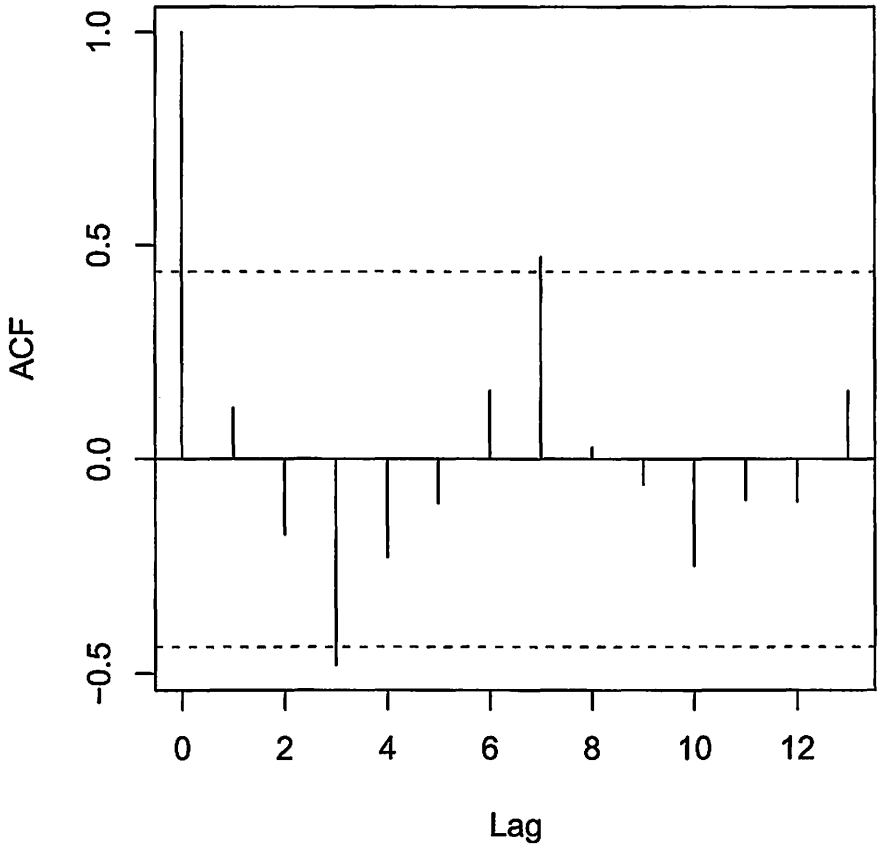

Figure 5: Autocorrelation plot of batch means of AR(1) time series.

```
        One Sample t-test

data:  batch
t = -1.4645, df = 19, p-value = 0.1594
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -2.6409967  0.4666197
sample estimates:
mean of x
-1.087188
```

2.2 Nonoverlapping Batch Means

```
> library(mcmc)
> bmvar <- olbm(out, batch.length = blen)
> length(out) * as.numeric(bmvar)

[1] 6437.297

> blen * var(batch)

[1] 5511.202
```

Note that olbm estimates σ^2/n, the asymptotic variance of the sample mean of the whole time series.

2.3 Initial Convex Sequence

Figure 6 (page 9) is made by the following R statements

```
> par(mar = c(5, 4, 1, 1) + 0.1)
> iout <- initseq(out)
> xx <- seq(along = iout$Gamma.con) - 1
> yy <- 1/(1 - rho^2) * (1 + rho) * rho^(2 * xx)
> ymax <- max(yy[1], iout$Gamma.con[1])
> plot(xx, iout$Gamma.con, type = "l", xlab = "Index (half lag)",
+     ylab = expression(tilde(Gamma)), ylim = c(0,
+         ymax))
> lines(xx, yy, lty = 3, lwd = 2)


> initseq(out)$var.con

[1] 7408.993
```

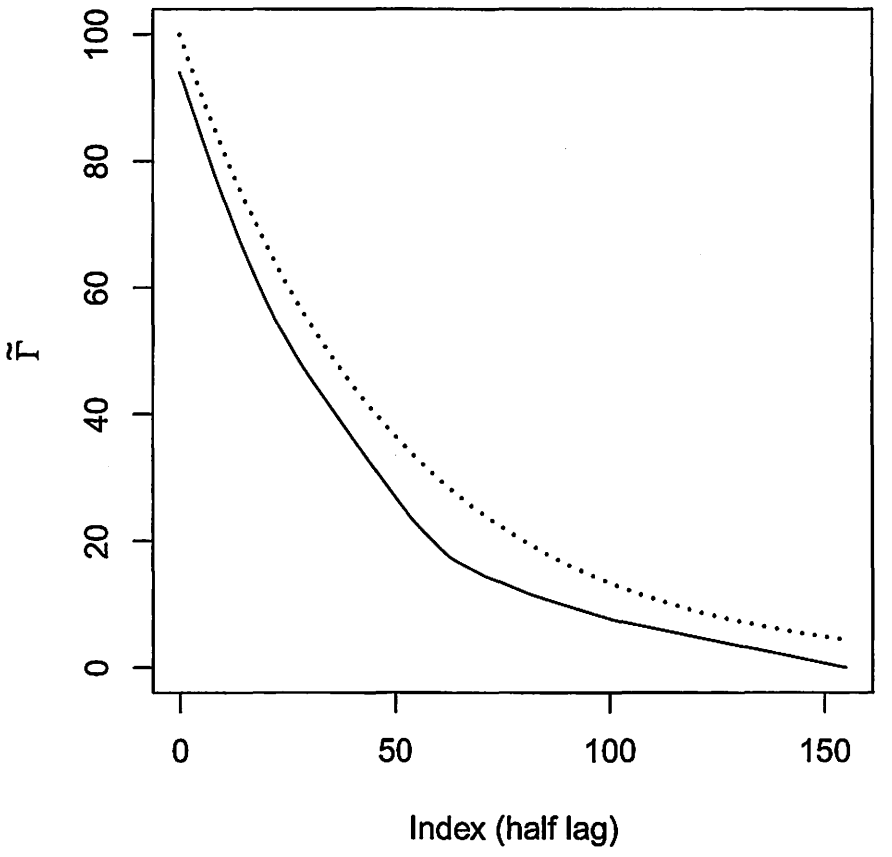

Figure 6: Autocovarince plot for AR(1) time series using initial convex sequence estimator.

2.4 Initial Convex Sequence and Batch Means

```
> blen <- 50
> batch <- matrix(out, nrow = blen)
> batch <- apply(batch, 2, mean)
> blen * var(batch)

[1] 1973.089

> blen * initseq(batch)$var.con

[1] 7488.409

> 1/(1 - rho^2) * (1 + rho)/(1 - rho)

[1] 10000
```

3 Burn-In

Simulate an AR(1) time series purporting to show a need for burn-in to be compared with Figure 1. We make all the simulation parameters the same as those for Figure 1 except we make the starting point 10 standard deviations of the equilibrium distribution from the mean.

```
> rho <- 0.99
> nsim <- 10000
> sd.equilib <- sqrt(1/(1 - rho^2))
> sd.equilib

[1] 7.088812

> x <- 10 * sd.equilib
> for (i in 2:nsim) x <- c(x, rho * x[length(x)] +
+     rnorm(1))
> out <- as.ts(x)
```

Figure 7 (page 11) is made by the following R statements

```
> par(mar = c(5, 4, 1, 1) + 0.1)
> plot(out, ylab = "x")
```

References

Brooks, S., Gelman, A., Jones, G. and Meng X.-L. (forthcoming). *Handbook of Markov Chain Monte Carlo*. Chapman & Hall/CRC Press.

GEYER, C. J. (2005). *R package mcmc (Markov chain Monte Carlo), version 0.5-1*. `http://www.stat.umn.edu/geyer/mcmc/`.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. `http://www.R-project.org`.

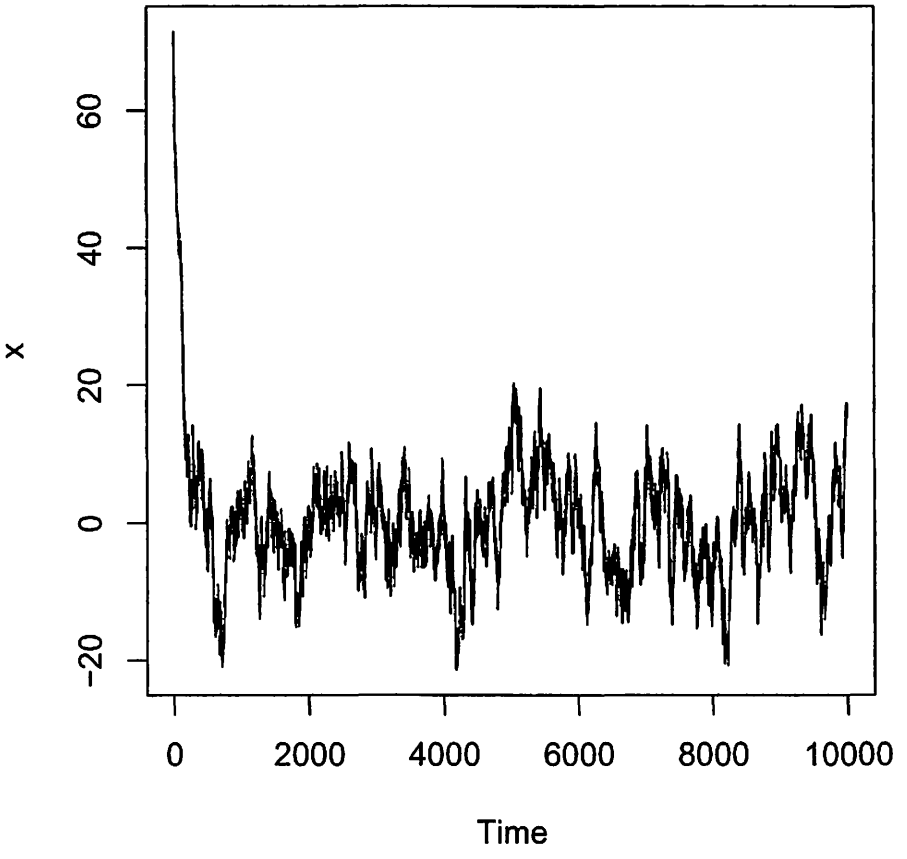

Figure 7: Time series plot of AR(1) time series. Compare Figure 1.